Cognitive-Based Metrics to Evaluate Collaboration Effectiveness

Dr. David Noble
Evidence Based Research, Inc.
1595 Spring Hill Road, Suite 250
Vienna, VA 22182-2228
USA
(703) 287-0312
noble@ebrinc.com

Dr. Michael Letsky
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217-5660
USA
(703) 696-4251
letskym@onr.navy.mil

ABSTRACT

Effective collaboration within culturally diverse multinational coalitions is essential in many military operations, especially in OOTW. Unfortunately, effective collaboration is sometimes difficult to achieve within any collaboration team. Because methods to improve collaboration, including selecting the right team members, creating the right type of organization, providing the right kind of training, and selecting the right types of collaboration tools are not fully understood, identifying effective interventions requires experimentation. Metrics, and especially cognitive oriented metrics that focus on team member understandings, are critical to such experimentation. Such cognitive-focused metrics can measure not only whether particular interventions are improving team effectiveness, but can also illuminate the cognitive reasons for the improvement.

This paper reports on a three-year research effort to develop, test, and apply such metrics. It describes a model-based strategy for selecting metrics, several models useful for metrics generation, eight classes of metrics for measuring collaboration effectiveness and the factors that contribute to this effectiveness, and the results of two metrics evaluations that demonstrated the practicality of applying the metrics in military experiments. The handling of human and organizational issues, scenario development, selection of metrics, and use of models followed the recommendations of the Code of Best Practice.

Key Words: *Metrics, collaboration, experimentation, models, cognitive.*

1.0 INTRODUCTION

Cognitive-based collaboration metrics measure how well each team member understands his missions, tasks, and teams, how well the team members work together, and how effective the team is in producing high quality timely products efficiently. Successful metrics will enable collaboration assessors to review what happens in a collaboration, and to understand the relationship between individual understandings, team behaviors, and team products. When collaboration and teamwork does not work well, a well-founded set of collaboration metrics can help pinpoint exactly where in the process a problem arose, and so can help suggest remedies to these problems.

This paper describes a set of proposed cognitive-based collaboration metrics. It describes the collaboration models that form the theoretical foundation for the metrics, describes the metrics themselves, and reviews the evaluation of the practicality and feasibility of the metrics in military experiments.

Paper presented at the RTO SAS Symposium on "Analysis of the Military Effectiveness of Future C2 Concepts and Systems", held at NC3A, The Hague, The Netherlands, 23-25 April 2002, and published in RTO-MP-117.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **00 DEC 2003**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**
4. TITLE AND SUBTITLE **Cognitive-Based Metrics to Evaluate Collaboration Effectiveness**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Evidence Based Research, Inc. 1595 Spring Hill Road, Suite 250 Vienna, VA 22182-2228 USA; Office of Naval Research 800 N. Quincy Street Arlington, VA 22217-5660 USA**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release, distribution unlimited**		
13. SUPPLEMENTARY NOTES **See also ADM001657., The original document contains color images.**		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**46**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

We define collaboration here to be "the mental aspects of joint problem solving for the purpose of achieving a shared understanding, making a decision, or creating a product." This definition emphasizes the cognitive and problem solving aspects of collaboration, as opposed to other definitions that place greater emphasis on information sharing. For example, the Information Superiority Working Group (Alberts, et al, 2001) defines collaboration as "actors actively sharing data, information, knowledge, perceptions, or concepts when they are working together toward a common purpose and how they might achieve that purpose efficiently or effectively."

2.0 THEORETICAL FOUNDATION – COLLABORATION MODELS

A cognitive-focused theory of collaboration describes the mechanisms that connect team member understandings to team effectiveness. It accounts for how the quality, completeness, and alignment of team members' understandings impact team performance and the quality of team products. The cognitive-based models of collaboration describe the theory.

Because there are many different factors to consider in understanding these connections, it is awkward to represent all of these factors in any single model. Instead, it is more practical to develop separate models that address different factors. These models do not present competing or conflicting interpretations of collaboration. Instead, they complement each other, each clarifying different aspects of collaboration. Five different collaboration models have contributed to development of the cognitive-focused collaboration metrics. These are the models describing teamwork and "taskwork," planning-execution feedback for both teamwork and taskwork, the interplay between teams whose members that oscillate between working separately and gathering as a team, the relationship between cognition, tasks, and products, and the importance of "transactive memory" as a key intervening variable. The first three of these models are described in the Phase 1 SBIR report (Noble, et al, 2000). The remainder are described in the metrics evaluation report (Noble, et al, 2001).

2.1 Teamwork/Taskwork Model

This model (Figure 1) describes the framework for organizing the different activities that teams must do. It distinguishes "teamwork" from "taskwork," terms that we have adopted from the UK CP-21 project. Taskwork is the work that the team must do to accomplish its mission, ignoring the coordination and other additional work that arises from working as a team. Teamwork is the additional work that the team must do in order to function as a team. It includes deciding how to partition the work among team members, how to coordinate their efforts, and how to adjust the team when necessary.

	Teamwork Organize and maintain the team	Taskwork Develop the product
Prepare	**Organize for teamwork** • Agree on goals • Identify tasks • Assign roles • Develop schedule • Identify interaction criteria and methods	**Develop mission plan** • Analyze mission • Identify tasks • Allocate tasks • Develop schedule • Assign resources • Identify constraints • Develop contingencies
Execute	**Attend to team health** • Monitor team processes • Cue and alert to possible problems • Diagnose nature of team problem • Reengage "organize for teamwork"	**Execute mission plan** • Monitor • Assess situation • Decide on needed plan adjustment • Issue directives • Execute / develop products

Figure 1: Teamwork/Taskwork Model.

This model enumerates and organizes many of the functions that teams must do. Because each of these functions can be important in collaboration under some circumstances, each requires metrics. For example, the model specifies that when teams organize for teamwork, they first need to agree on goals. Therefore, the model implies a need for metrics on the extent that people agree on goals and on the efficiency of the processes by which they reach agreement.

2.2 Planning/Execution Feedback Model

This model (Figure 2) builds on the teamwork/taskwork model. It emphasizes the importance of monitoring and adjustment for both teamwork and taskwork. In the case of taskwork, this corresponds to the normal feedback that occurs during plan execution. Commanders monitor the progress of a plan to determine if the plan will still enable them to achieve their objectives. If it does, they continue to execute the plan. If it does not, then they adjust the plan so that it will. Similarly, when a team is working together, the members need to monitor the team organization and processes, and to make adjustments to the team when needed. Common adjustments are to provide additional help to team members' who are overloaded, to supplement the expertise of the team, or to reassign roles to leverage team members' skills better.

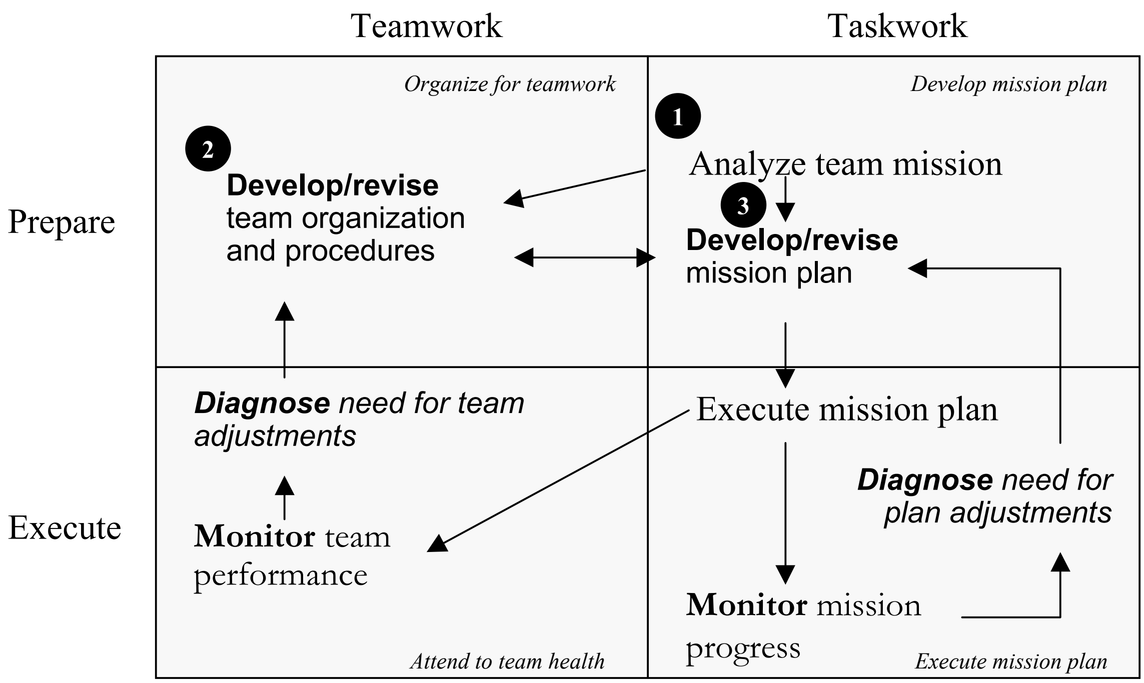


Figure 2: Planning/Execution Feedback Model.

This model implies the need for metrics on how well teams monitor their mission and team performance. Metrics for taskwork – the mission performance – have long been recognized as important in measuring command and control effectiveness (Hayes et al, 1983). This model emphasizes the need for an analogous set for measuring the feedback processes concerned with team health. For example, it suggests such metrics as time to detect that a team member is overworked and requires backup.

2.3 Individual-Team Interplay Model

The framework in the previous two models applies to every collaborative team. The individual-team interplay model (Figure 3) applies to only some types of teams. It describes the interactions important in teams such as collaborative planning groups. In these teams, team members occasionally meet synchronously as a full team to discuss and resolve issues and to adjust individual tasks. After this meeting, the individual team members separate to work on their separate tasks.

When meeting synchronously, the group exchanges information about team and task issues to develop a consensus. Figure 3 lists several different types of information exchange that frequently occur. These include distributing information, discovering differences, brainstorming, critiquing and enriching each other's ideas, guiding, and negotiating, and making decisions.

After the meeting, team members separate to work on their individual tasks (left side of Figure 3). In performing their tasks, they continually make decisions about what they should do. Figure 3 represents this individual behavior by listing seven cognitive functions important to decision making. Though not meeting synchronously as a whole team, team members do not work in isolation. They interact by sharing documents and other computer products (visualizations) and by talking with each other. Occasionally, one or more team members will decide that the whole team needs to meet again. Team members discuss an agenda and then separately prepare for the meeting.

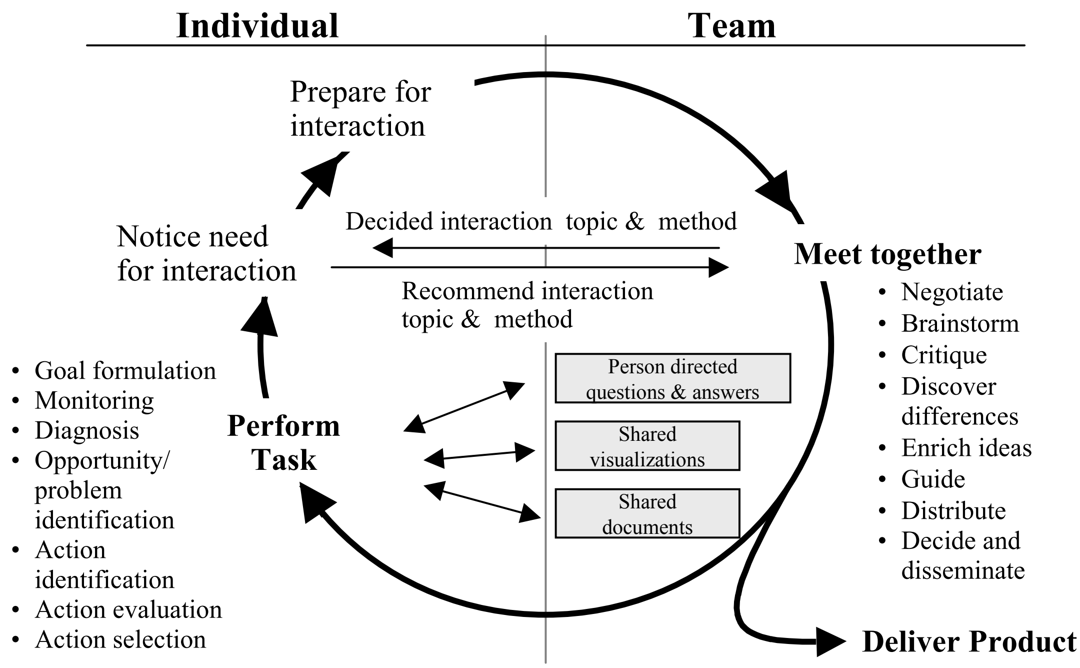


Figure 3: Individual-Team Interplay Model.

This cycle of team meeting and individual task performance continues as each team member develops and assembles the parts of the product he is responsible for. This process ends when the team delivers its final product.

This model lists the principal individual and group cognitive processes, each of which will sometimes be important to measure. The first EBR report on collaboration (Noble et al 2000) listed metrics for each of the cognitive steps in decision making (e.g., goal formulation, monitoring…), both for taskwork and for teamwork. The model also identifies some of the key synchronous team meeting processes that need to be measured. For instance, a metric for discovering differences could be the number of inconsistencies between people's understandings that nobody on the team is aware of.

2.4 Coupling Cognition, Behavior and Products

The Cognition-Behavior-Product model (Figure 4) emphasizes the nature of the relationship between individual and team understandings, individual and team behaviors, and individual and team products. It makes three important contributions: that team understandings, behaviors, and products must be mediated by individuals, that task quality and understandings affect each other, and that it is critical to measure individual task performance to assess collaboration effectiveness.

This model emphasizes that team cognition, team behaviors (the items listed under "meet together" in the individual-team interplay model) and team products connect only through individual efforts. For example, information exchange may be needed when different team members interpret team goals differently (differences in understanding are part of team understanding) and are aligned by discovering differences, clarifying ideas, and negotiation (team behaviors). However, this clarification only occurs when one or more team members individually realize that there is a cognitive difference that's important to address, and when each individual engages in the behavior necessary to address it.

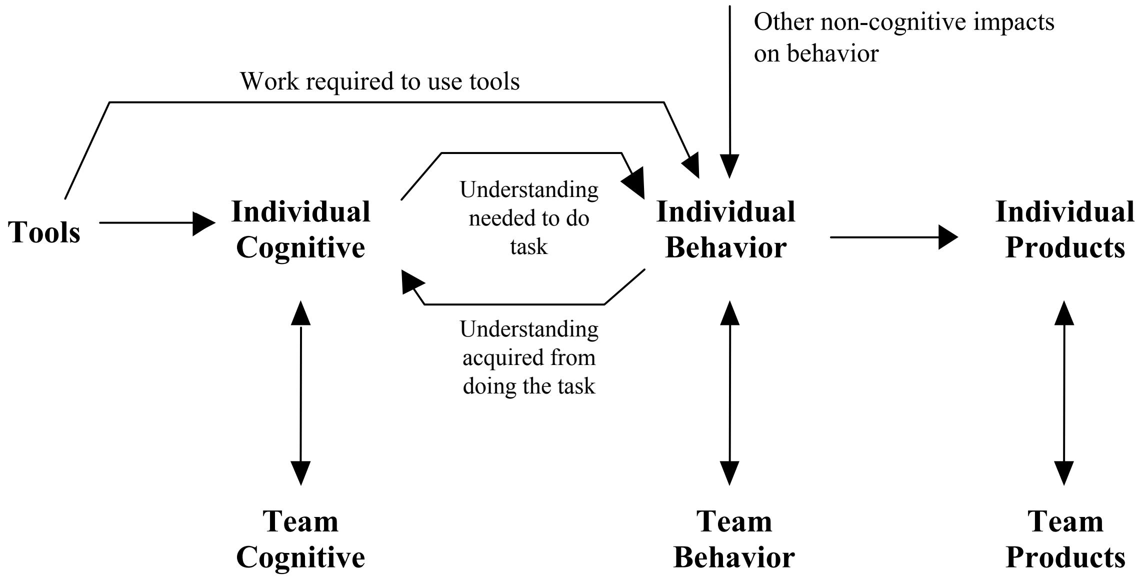


Figure 4: Cognition-Behavior-Product Model.

The second important feature of this model is the bi-directional relationship between an individual team member's understandings (individual cognitive) and his task performance (individual behavior). The forward direction, from understanding to task, is obvious. People who did not understand a task well enough to perform a task well will usually perform poorly. The backward direction is less obvious, but very important to understanding the dynamics of collaboration. That is, people who don't perform a task adequately will fail to acquire the understanding that doing the task well would provide. If that understanding is needed to support subsequent related tasks, the team members would then also fail to perform those tasks well, which in turn would undermine performance on additional tasks. Hence, failure to understand what's needed in order to perform an early task well can set up a chain reaction that undermines a long sequence of additional dependent tasks.

This model points out that individual task performance mediates understandings and product development. Accordingly, it stresses the importance of measuring task performance, to include how well a task is accomplished and adherence to the task schedule.

2.5 Transactive Memory Model

The transactive memory model (Figure 5) for collaboration has been developed and tested over the past fifteen years by team researchers, Moreland, Argote and Ingram, from the University of Pittsburgh, Carnegie Mellon University and Columbia University, respectively (Liang et al, 1995; Moreland and Myaskovsky, 2000; Argote and Ingram, 2000). Because of its emphasis on individual and team cognition and its strong empirical foundation, this model has been especially useful in identifying powerful collaboration metrics.

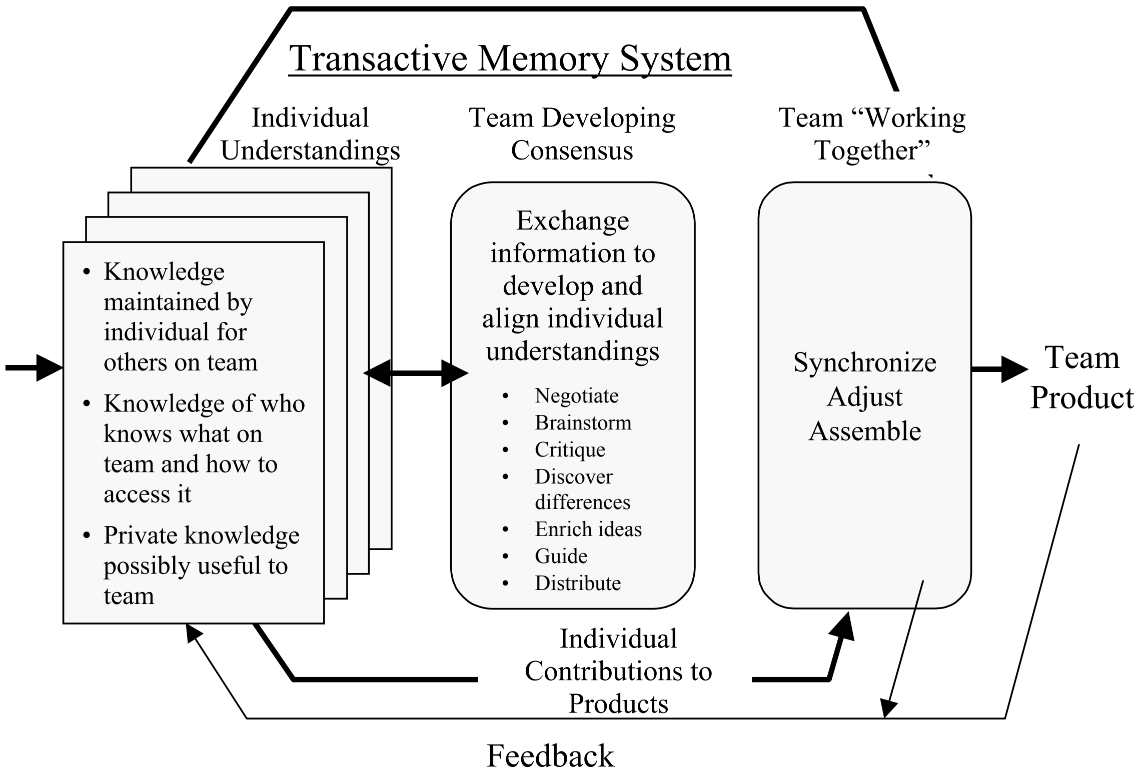


Figure 5: Transactive Memory Model.

The transactive memory itself consists of the collection of individual understandings and the team mechanisms to exchange information and so update these individual understandings. The individual understandings include all of the understandings about teamwork and taskwork pointed out in the teamwork/taskwork model. These include understandings about how to do the tasks required to perform the mission, understandings about the status of the situation and task, understandings of how the team is organized to function, and understandings about how the team is actually functioning now. It includes the common ground elements such as understanding of other team member's capabilities, workload, and knowledge.

In the transactive memory model, every team member is not expected to know everything. Instead, the knowledge is distributed throughout the team. As indicated in Figure 5, the model classifies individual understandings in terms of their relationship to the understandings of other team members. This classification emphasizes how team members leverage each other's knowledge. The classes shown in Figure 5 are the knowledge that individual maintains for team, knowledge about what others maintain for the team and how access that knowledge, and private knowledge that the individual should share with others if that knowledge becomes relevant to the team. Team members also have a "meta-knowledge," an assessment of the adequacy of their knowledge. In addition, they have an assessment of what they believe the team as a whole has decided.

The "team developing consensus" block includes the same elements as the "meet together" function in the "individual-team interplay model." In this model, however, the purpose and effects of these interactions are to update and align the individual understandings in the transactive memory system.

Transactive memory appears to be a powerful intervening variable for collaboration. The transactive memory literature has shown that in those cases examined (e.g., team training methods) the state of the transactive memory can account for (statistically) all of the effects of the experimental manipulation (Liang et al, 1995).

If this finding should generalize, then the transactive memory suggests a powerful class of cognitive collaboration metrics. This is the completeness and accuracy of a team's transactive memory, compared to what the transactive memory needs to be in order for the team to interact effectively. If, as the model suggests, the purpose of team member information exchanges are to create this needed transactive memory state, then the effectiveness of these exchanges may be assessed in terms of their impact on that state.

This last point, that the primary purpose of information exchange within a team is creation of the transactive memory that a team needs to carry out its tasking effectively, has great significance to metrics for collaboration. It suggests that measuring the amount or type of communication that occurs in a team is not particularly useful for understanding teamwork, unless that measurement can be related to its impact on the team's transactive memory.

3.0 COGNITIVE-FOCUSED COLLABORATION METRICS

Because teams collaborate for many different reasons and work together in so many different ways, there are a large number of potentially useful metrics. Figure 6 organizes these diverse metrics into four individual and team categories implied by the Cognition-Behavior-Product and Transactive Memory models. The following material describes the metrics in each of these categories, starting with the products and working toward understandings.

	Understandings	**Information Interactions**	**Task Performance**	**Products**
Individuals	Individual Understandings	Individual Information Interaction Support	Individual Task Performance	Individual Products
Teams	Team Understandings	Group Information Interactions	Teamwork	Team Products

Figure 6: Categories of Collaboration Metrics.

3.1 Product Metrics

These metrics measure product quality and timeliness, and the efficiency with which they are produced. They are the "proof of the pudding" metrics, for teams normally should not be designated to be effective unless they can produce a good product efficiently.

Metrics for products produced by teams are the same as those produced by individuals, for the measurement of product quality should not depend on how the product was produced. In command and control, typical products are situation assessment briefings, plans, and decisions. The quality of many of these, like plans and decisions, are difficult to measure because there is usually not a known "book solution" that can serve as a standard for correctness and completeness. Though measuring such abstractions is challenging, EBR has been doing so successfully for more than a decade. HEAT (Hayes et al 1983) describes many of these metrics.

Examples of product measurements are the quality and timeliness or a situation assessment briefing and the efficiency with which it's produced. Timeliness is product creation time relative to a deadline. Team efficiency is the total person hours required to complete a product.

Situation assessment quality has been measured by comparing team member's assessments to the assessments of experts made under ideal conditions. The metrics measure the correctness of team member's situation assessments in such categories as identity, location, and capabilities of own, neutral, and adversary forces; adversary and own opportunities and risks; adversary intentions and possible courses of action; and key environmental factors.

3.2 Task Performance Metrics

These metrics measure the processes for creating and assembling products. When applied to individual team members, they measure task performance, schedule adherence, workload, level of engagement, and flexibility. When applied to the team, they measure how well the team synchronizes, adjusts, and assembles its products. These metrics can be highly diagnostic of overall team effectiveness, with significant impact on the team product quality and efficiency metrics. Tables 1 and 2 summarize these metrics applied to individuals and teams.

Table 1: Metrics for Individual Task Performance

Issue Metric Addresses	Metrics
Overall performance	Fraction of tasks not addressed Thoroughness with which a task is done Correctness of task process employed
Schedule adherence	Number of tasks completed early and late Amount of delay in start time If completed late, how late Number of tasks out of order
Workload	Fraction of time team member is idle Fraction of assigned work not completed when no idle time
Level of engagement	Fraction of time team member devotes to task
Flexibility	Fraction time schedule is adjusted when needed Fraction time type of task is adjusted when required Fraction of time nature of task product is changed when required

Table 2: Metrics for Team Performance Emergent Behaviors

Issue Metric Addresses	Metrics
Team agility	Time required to note that team needs to adjust Fraction of time that adjustment is attempted when needed Effectiveness of adjustment
Synchronization	Average delay in starting a task because precursor tasks were delayed Diminishment of desired effect because of imperfect synchronization
“Fibrillation”	Fraction of preliminary individual products never used Fraction of individual products needing revision before they can be used

There are two categories of task performance metrics at the team level. The first type of metric aggregates the individual task performance metrics. It might, for example, take the average of team members’ workloads. The second type measures emergent team behaviors. These are behaviors that apply to the team as a whole, but cannot be defined at an individual level. An example is “fibrillation,” in which there is a substantial amount of work being done by individual team members, but the work does not contribute in a coherent way to an overall product.

3.3 Information Interaction Metrics

These metrics measure the adequacy of brainstorming, negotiating and the other processes that the team employs to acquire the required shared understandings and team consensus. At the individual level, these metrics measure individual contributions in support of developing the group understandings and consensus. At the team level, they measure the effectiveness of various group processes themselves.

The individual level metrics focus on the effectiveness of transfer of meaning, on the extent that each team member acquires the right information from an appropriate source, and on the extent the he provides the needed information to the appropriate recipient. Table 3 summarizes several of these metrics.

Table 3: Metrics for Individual Team Member Information Interactions

Issue Metric Addresses	Metrics
Information acquisition	Fraction of time correct team member is asked for information
Information provision	Fraction of time “private information” needed by group is provided
Transfer of meaning	Fraction of time information needed by others is provided in a way that could be understood without the need for clarification

The team level information interactions address how well the team as a whole functions as an assessment and decision making entity. An effective team will identify good lists of candidate assessments and actions and will evaluate these lists considering a full range of relevant criteria. Good teams will avoid the information filtering and biased evaluation criteria documented to arise in “Group Think” (Janis, 1972). Table 4 provides examples of metrics in this category.

Table 4: Metrics for Team Information Interactions

Issue Metric Addresses	Metrics
Brainstorming	Completeness of alternative situation interpretations considered by group Complete of decision alternatives considered by group Completeness of decision criteria considered by group
Negotiating	Fraction of time people advocating conflicting actions find an action acceptable to all parties
Critiquing and idea enrichment	Fraction of people on team responsible for an area asked to comment on products in that area Fraction of time spent in a meeting not relevant to own responsibilities and not contributing to other
Discovering differences	Fraction of differences in understanding identified
Distributing	Fraction of the people who should receive information that actually receive the information Average fraction of irrelevant information received by each team member

3.4 Cognitive Metrics

Cognitive metrics measure the extent to which the team understands what it needs to understand in order to be effective. That is, they measure the adequacy of a team's transactive memory system.

The EBR report, "Metrics for Evaluation of Cognitive Architecture-Based Collaboration Tools, (Noble 2000) identified hundred of cognitive metrics for individual team members. These are organized in terms of teamwork and taskwork for seven decision making processes: goal formulation, monitoring, situation diagnosis, opportunity/problem identification, identification of candidate actions, evaluation of these candidates, and action selection. They address all of the elements of transactive memory, such as knowing what team member possesses needed team knowledge and knowing how to access that knowledge. Table 5 lists a few of these metrics.

Table 5: Cognitive Metrics for Individual Team Members

Issue Metric Addresses	Metrics
Taskwork: Understanding commander's intent	Correctness of team member's understanding of commander's intent
Taskwork: Situation understanding	Correctness of team member's estimate of adversary goals
Taskwork: Schedule and process information	Correctness of knowledge of deadlines
Teamwork: Knowing team member responsible for various kinds of knowledge	Correctness and completeness of knowledge each team member is responsible for
Teamwork: Identify team member overworked and likely to not finish task on time	Correctness of team member's estimate of workload for those team members producing a product needed as input to that team member

There are three different types of team level cognitive metrics: roll-ups, team coverage, and alignment. Roll-ups are averages of the team member's metrics, averaged over team members. Team coverage concerns gaps in team knowledge or maximum expertise of knowledge within a team. Alignment concerns the extent that team members' understandings are consistent.

4.0 EXPERIMENTAL DEMONSTRATION OF METRICS PRACTICALITY

The metrics identified in the previous section must be feasible to be useful. That is, it must be practical to collect the data to compute the metrics in the experiment environments in which it is desired to measure collaboration effectiveness. EBR performed two evaluations of the above metrics, piggy backing on experiments at PACOM at Camp Smith, Hawaii and at JFCOM at Suffolk, Virginia. Both of these experiments confirmed the feasibility of the metrics.

4.1 Evaluation Issues

The practicality of these cognitive-focused collaboration metrics may be limited by their large number and by collection constraints at operational venues. In addition, use of some of these metrics may be hindered by the low observability of the phenomena being measured; by the large amount of data needed, and by their high level of abstraction. Because of these potential problems, the metrics evaluation sought to answer the following questions.

1) Does insight about collaboration require so many metrics that collecting the needed data to estimate this number is impossible?
2) Do the data collection constraints during experiments at military sites preclude obtaining the required data?
3) Does the low observability of cognitive metrics (e.g., measuring what people know) preclude collecting the needed data?
4) Is it possible to collect the volume of data needed to compute team level metrics which require measurements of all team members?
5) Are the team product metrics developed to measure C2 processes fully applicable to measuring collaboration products?

At the two evaluations, EBR tried out the metrics determine the extent that each of the potential problems impacts their utility in actual experiments. The ACOA MUA evaluated the first four of these metrics feasibility questions, providing affirmative answers in each case. The JFCOM experiment addressed all of the issues under more stringent data collection conditions, and also provided affirmative answers. In both cases, EBR's data collection goals were added to the objectives of larger evaluations previously planned for other purposes.

4.2 Evaluation at ACOA Military Utility Assessment (MUA)

The ACOA (Adaptive Course of Action) Military Utility Assessment is a formal evaluation event required for any technology developed as part of Advanced Concept Technology Demonstrations (ACTD). ACOA is a suite of integrated tools to support distributed collaborative planning at the CINC and JTF levels.

At the ACOA Military Utility Assessment a group of military and government civilian planners evaluated the effectiveness of an advanced distributed planning system being developed by the Adaptive Course of Action ACTD. These planners manned spatially distributed workstations that provided access to the ACOA collaboration tools.

Data collection. A contractor working as part of this ACTD planned and conducted this evaluation, and analyzed the results. EBR was invited to contribute two data collectors who were assigned to two of the workstations, and to contribute questions that were included in questionnaires presented to each participant at eight designated times in the two day evaluation.

EBR was limited to 2 to 5 cognitively focused questions at each of these eight prescribed points in the evaluation. As part of the data collection constraints, EBR designed these questions so that each could be answered within a few seconds. These questions probed the participants' understandings about issues important to task and team understanding. The taskwork questions asked about commander's intent, adversary goals, and plan elements and weaknesses. The teamwork questions asked about the responsibilities of people at the different workstations, how busy they are, and whether the team needed additional outside expertise.

Unlike the data collected through the questionnaires, which were presented on eight occasions to all participants, the metrics-related data was collected continuously but at only two workstations. These data focused on observable behaviors and conversations, but also included participant comments on current concerns and issues.

Results. Asking participants questions about their taskwork and team understandings was sufficient to compute both the individual and team level cognitive metrics on these subjects. While limited to only a few topics, the completeness and accuracy of their answers, as computed by comparison with an answer key, provided insight about their level of teamwork and taskwork understandings in general.

The observer notes on participant behaviors, conversations, and comments were sufficient to understand the relationships between task performance and their cognitive and non-cognitive causes. In fact, it was the analysis of these data that gave rise to the Cognition-Behavior-Product Model depicted in Figure 4. Table 6 summarizes these relationships for the evaluation participant playing the role of the operational planning team leader. Note that in ACOA confusion over how to use these new kinds of tools added to operator workload on several occasions.

Table 6: Summary of Impacted Behaviors and their Causes for the MUA Team Leader

		Cognitive Causes					Non-cognitive Causes			
	Total Impacts	Tool mechanics confusion	Tool concept confusion	Task/process confusion	Low awareness of other's activities	Judged not worth the effort	Machine requires extra effort	Waiting	Synchronization	Artificiality of environment
Task Impacts										
Low task quality	2		1							1
Not completed	5	1	1	2		1				
Cursory	3					3				
Delayed	4	1					2		1	
Out of order	2			2						
Increased workload	24	11	1	2	2		7			
Activity level										
low										
Engagement level										
low										
Total of Causes	41	14	3	6	2	4	9	0	1	1

4.3 Evaluation of JFCOM Presentation LOE

This second metrics evaluation was intended to be a more stringent test of the collaboration metrics feasibility. Like the ACOA MUA, the EBR collaboration metrics evaluation team was permitted to add questions to a questionnaire presented to participants every few hours. However, unlike the ACOA MUA, where EBR data collectors sat next to key participants and were free to ask questions throughout the experiment, the JFCOM experiment imposed the more typical experimentation constraints where observers were to be "flies on the wall" during the scenario execution. As described below, these constraints did not reduce the ability to collect cognitive data from questionnaires. However, they did reduce the ability collect behavioral data.

The JFCOM Limited Objective Experiments (LOE), held at JFCOM/JTASC in August 2001, compared the effectiveness of three alternative methods of presenting and interacting with situation information. In this experiment, 18 staff members from JFCOM were organized into three groups, each functioning as a collaboration team. Each member of each of these groups was assigned to one of six positions, and retained that position throughout the two weeks of the experiment. The positions were "chief of staff," "plans," "operations," "future information," "current information, and logistics." Each team worked together in a room dedicated to a particular presentation method, but members were separated by partitions to generate the effects of spatially distributed teams. Team members shared a large wall-mounted visualization and personal computer visualizations. They could communicate by voice or by e-mail.

Data collection. The experiment exposed each of the three groups to each of the alternative presentation methods. The experiment was divided into twelve time periods. After each period, each participant answered a questionnaire about the situation, and each of the three groups prepared a situation briefing. In the data analysis, the experiment analysts scored each individual questionnaire for correctness and completeness of situation understanding, using an answer key that represents expert understanding. The analysts also scored the team situation briefing using the same answer key.

Because the JFCOM experiment focused on the effectiveness of information presentations in supporting situation understanding, many of the experiment questions prepared by the experiment organizers were the same as those needed to test the cognitive metrics for taskwork.

Results. As in the ACOA experiment, this experiment demonstrated the feasibility of the metrics. The desired data were able to be collected and analyzed under the fairly restrictive constraints placed on the data collectors. This experiment also showed the applicability of product quality metrics to the products produced by teams.

Figure 6 portrays a particularly significant result from the JFCOM experiment – the strong confirmation for the substantial advantages of collaboration, with the data being consistent with the hypothesis that these advantages are mediated by the "transactive memory mechanism" discussed in Section 2.5. This figure compares for each of the three groups in the experiments the 1) the average situation understanding among team members, 2) the best individual situation understanding within the team, and 3) the quality of the team's briefing describing the situation. As shown in Figure 6, for each group the team briefing was significantly better than the average situation understanding of team members, and was in fact significantly better than the best understanding of any individual team member. Note that the criteria for evaluating the briefings and the situation understandings were identical. Both were based on the same answer key and on the same scoring criteria.

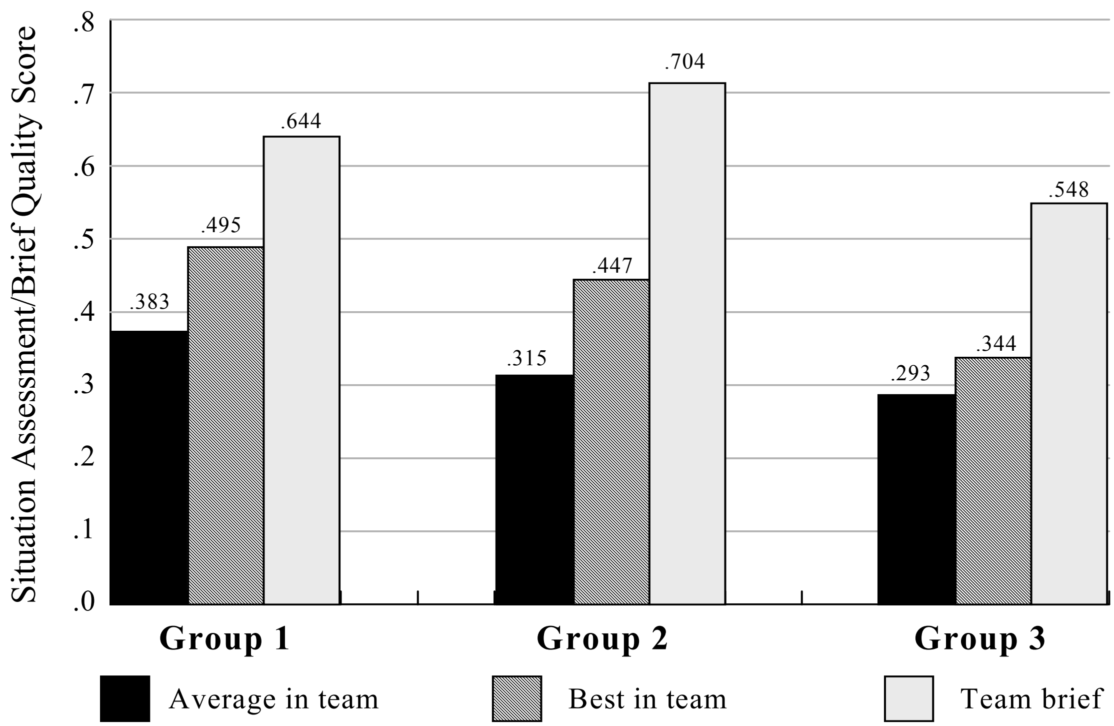


Figure 6: Measured Quality of Team Situation Assessments and Briefings.

Without looking at the data for "best in team," the improvement of the team brief over the average understanding had two plausible explanations. The first is that the best member did the brief for the rest of the team, not relying on or needing the input from the other team members. The second is that the team members pooled their individual understandings, with each team member contributing especially to his particular area of understanding.

Clearly, in this JFCOM experiment, the "best in team" data rules out the first explanation, because for every team, the team brief was significantly better than the best of the individual situation understandings.

This example shows not only that team cognitive metrics are feasible to collect and compute, but that these metrics support theory development and testing. These metrics clearly support the transactive memory model, for this model provides a very direct way to understand these results. This model asserts that a collaborating team divides up responsibility for knowing various facts and procedures among team members. Each team member knows who is responsible for a particular area and knows how to obtain that information from those team members.

The transactive memory mechanism can easily explain the results observed in this experiment. Because someone knows something about each subject, pooling team knowledge generates some information about each aspect of the situation. This generates the high briefing score. Because no one knows everything about the situation, the understanding of any one individual, including the best informed person, is less complete than that of the team as a whole.

5.0 CONCLUSION: APPLICATION OF CODE OF BEST PRACTICE

The cognitive-focused collaboration metrics measure individual and team understandings, information interactions, behaviors, and products. They are feasible to employ, and can not only measure team effectiveness but can provide insight into the reasons for effectiveness.

Development and evaluation of the cognitive metrics complied with the recommendations of the Code of Best Practice:

- *Metrics organization.* They are organized into a hierarchy of metrics types that address the quality of the overall team product as well as the effectiveness of the understandings and behaviors the contributed to developing the product.
- *Human factors*. They address an important human factors issue – the relationship between individual understandings and behaviors and overall team effectiveness.
- *Scenarios.* The scenarios used to evaluate the metrics were designed to exercise key factors expected to drive utility of the metrics.
- *Use of models*. They are informed, guided, and motivated by cognitive models of collaboration.

6.0 REFERENCES

[1] Alberts, David; Garstka, John; Hayes, Richard; and Signori, David; "Understanding Information Age Warfare", Command and Control Research Program, 2001.

[2] Argote, Linda; and Ingram, Paul; "Knowledge Transfer: A Basis for Competitive Advantage in Firms", *Organizational Behavior and Human Decision Processes,* Vol. 82 No. 1 (May 2000): 150-169.

[3] Hayes, Richard E.; Hainline, Mark; Strack, Conrad; and Bucioni, Daniel; "Theater Headquarters Effectiveness: Its Measurement and Relationship to Size, Structure, Functions, and Linkage", McLean, VA: Defense Systems, Inc., 1983.

[4] Janis, Irving; "Victims of Groupthink", Houghton Mifflin Company, 1972.

[5] Liang, Diane Wei; Moreland, Richard; and Argote, Linda; "Group Versus Individual Training and group Performance: The Mediating Role of Transactive Memory", *Personality and Social Psychology Bulletin*, Vol. 21 No.4 (April 1995): 384-393.

[6] Moreland, Richard; and Myaskovsky, Larissa; "Exploring the Performance Benefits of Group Training: Transactive Memory or Communication?", *Organizational Behavior and Human Decision Processes,* Vol. 82 No. 4 (May 2000): 117-133.

[7] Noble, David; and Buck, Diana; "Metrics for Evaluation of Cognitive Architecture-Based Collaboration Tools", Phase 1 SBIR Final Report, Evidence Based Research, Inc., 2000.

AUTHOR BIOGRAPHY

Dr. David Noble is a Senior Scientist at Evidence Based Research. He is currently conducting research on the cognitive foundations of collaboration, with the goal of developing theory-based guidelines for improving the effectiveness of collaboration teams. In the past Dr. Noble has conducted research on the cognitive-basis of situation assessment and decision making and on ways of measuring situation assessment and decision making performance. In addition, he has developed data fusion algorithms for decision aids to improve situation assessment and decision making. Dr. Noble received his doctorate in applied mathematics from Cornell University.

This page has been deliberately left blank

Page intentionnellement blanche

Cognitive-Based Metrics to Evaluate Collaboration Effectiveness

Presented at

SAS 039 Symposium

Analysis of Military Effectiveness of Future C2 Concepts and Systems
Den Hague, Netherlands
April 23-25

Presented by

Dr. David Noble, Evidence Based Research
Dr. Michael Letsky, Office of Naval Research

Evidence Based Research, Inc.
1595 Spring Hill Road, Suite 250
Vienna, Virginia 22182-6800
ebrinc@ebrinc.com

Office of Naval Research
Code 34
800 North Quincy Street
Arlington, Virginia 22182-6800
xxxc@onr..navy.mil

Topics

- Definitions
- Models
- Metrics
- Experiments
- COBP Application

Definitions

- Office of Naval Research Collaboration and Knowledge Management Thrust
  - the mental aspects of joint problem solving for the purpose of achieving a shared understanding, making a decision, or creating a product
- Information Superiority Working Group
  - actors actively sharing data, information, knowledge, perceptions, or concepts when they are working together toward a common purpose and how they might achieve that purpose efficiently or effectively

Metrics Strategy

- Identify relevant models of how collaboration works
- Use models to identify candidate metrics
- Conduct experiments to test feasibility and utility of metrics

Model Application

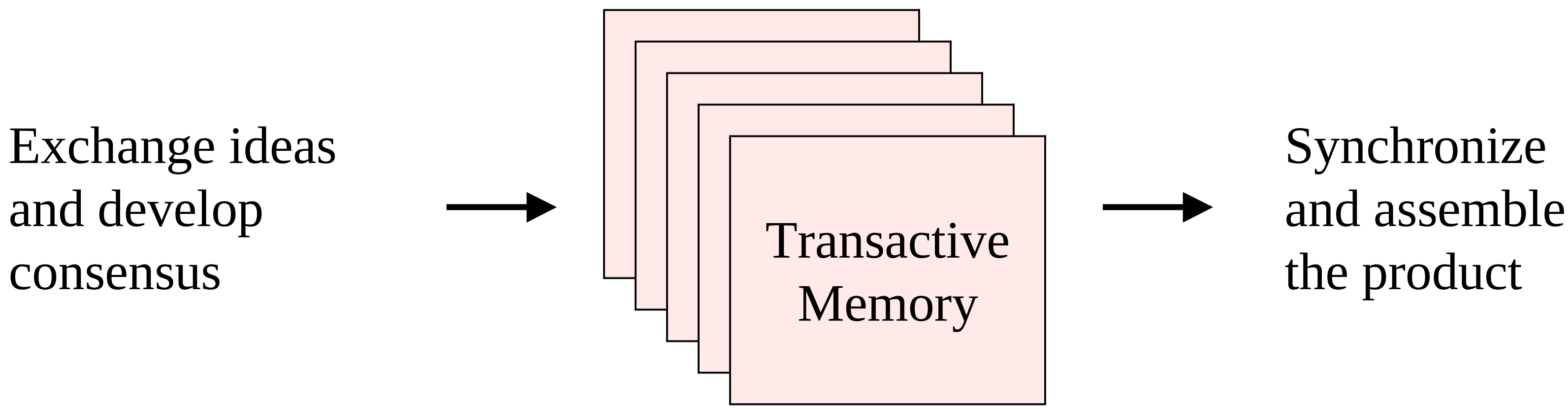
Individual Understandings	Individual Information Interaction Support	Individual Task Performance	Individual Products
Team Understandings	Group Information Interactions	Teamwork	Team Products

Models can help identify key variables to measure -- those that illuminate the connection between individual understandings to the quality of team products

Need for Complementary Collaboration Models

- No model can likely address all aspects of collaboration
- However, a set of complementary models, each addressing different aspects of the same underlying process, can
- Five models of value to metrics development
  - **Teamwork and Taskwork**: all teams engage in activities to develop tasked product and to maintain team health
  - **Feedback:** teams must monitor progress and make corrections for both teamwork and taskwork
  - **Individual/Team Interplay**: many collaborative tasks call for a cycle of individual and group processes
  - **Coupling Cognition, Behavior, and Products**: team processes and products emerge from individual understandings, behaviors, and products
  - **Transactive Memory**: the relationship among team member's individual understandings drives the quality of team performance

Transactive Memory As a Key Intervening Variable



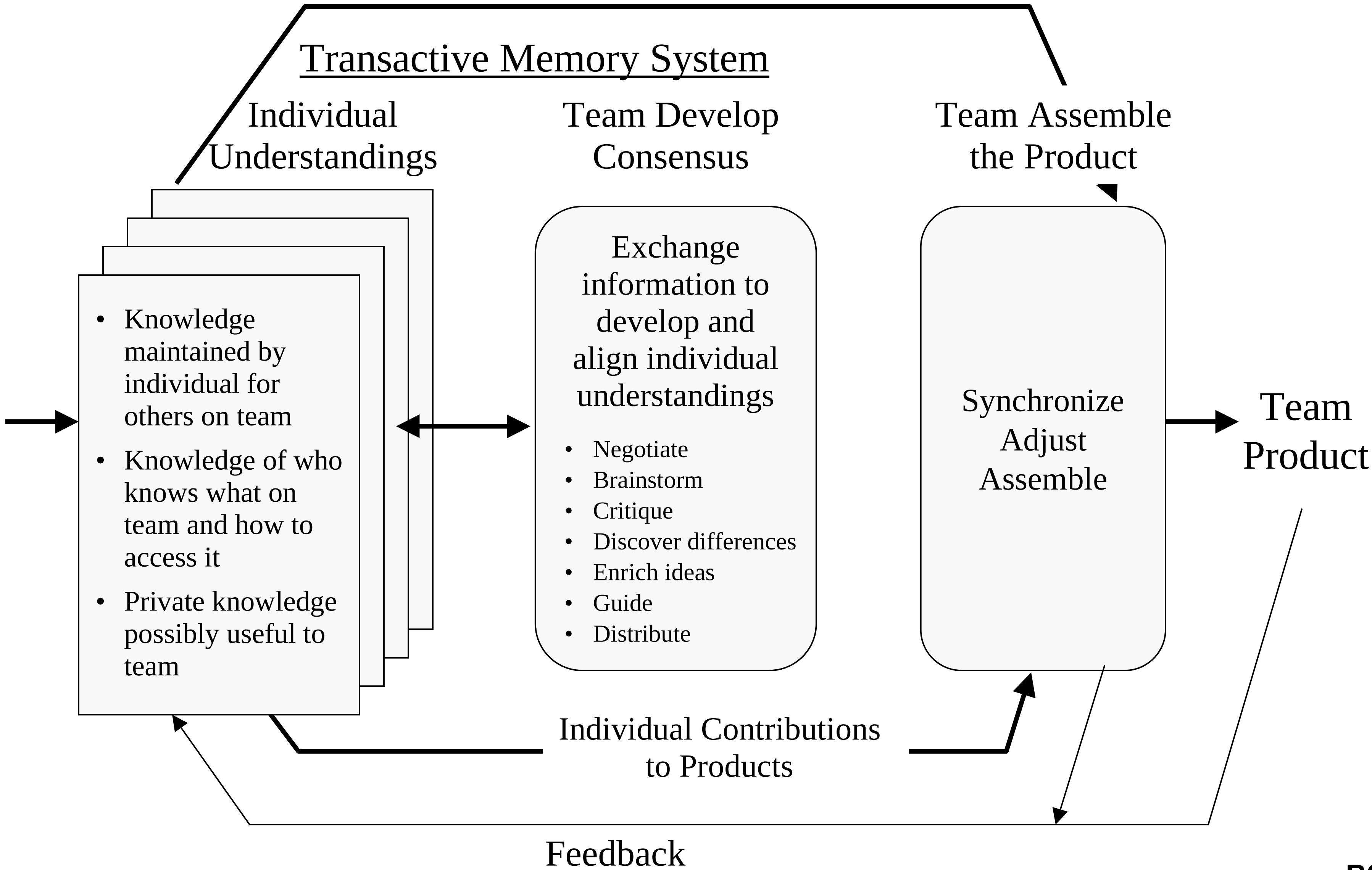
Transactive memory

describes the distribution of knowledge within a team

functions as a powerful intervening variable between group discussions and group behaviors

emphasizes importance of distribution of knowledge within the team and each person's understanding of how to obtain that knowledge

Transactive Memory Model



Metrics Categories

* Individual Understandings	* Individual Information Interaction Support	Individual Task Performance	* Individual Products
* Team Understandings	Group Information Interactions	* Team Behavior	* Team Products

* Selected for illustration in this brief

Metrics
Individual and Team Products

- Measures quality of products and efficiency with which they are produced
- Bottom line “proof of the pudding” metrics
- Metrics for a particular product are the same, whether produced by an individual or a team
- Situation Assessment example
  – Accuracy and completeness of user statements about the elements of the situation, compared to a situation answer key
  – Elements of situation assessment: adversary and own opportunities and risks, adversary capabilities and intentions, key environmental factors

Team Behavior Metrics

- Measures “mechanics” of teamwork (e.g., team agility, synchronization, “fibrillation” and “friction”)
- Examples
  - Time required for team to recognize a problem in teamwork or product development (agility)
  - Fraction of tasks delayed because needed precursors were late (synchronization)
  - Fraction of preliminary individual products never used (“fibrillation”)
  - Occurrence of individual products that cannot be assembled because they are incompatible (“friction”)

Individual Information Support Metrics

- – Measures individual performance to support development of group consensus
- – Includes information acquisition, formulation, and dissemination
- – Examples
  - Fraction of times right person asked for information
  - Fraction of time information needed by others conveyed in ways that could be understood without need for clarification

Individual Cognitive Metrics

- Team member understanding of status and processes for teamwork and taskwork
- Organized into decision process categories
  - Goal formulation, monitoring, situation diagnosis, opportunity/problem identification, identification and evaluation of candidate actions, action selection
- Examples
  - Correctness of team member understanding of commander's intent
  - Correctness of knowledge of task deadlines
  - Correctness of knowing knowledge that each team member is responsible for
  - Completeness of knowledge of situation elements relevant to that member's tasks

Team Cognitive Metrics

- Types
  - Roll-ups average individual cognitive metrics
  - Team coverage measures best knowledge in team and gaps
  - Alignments summarize extent of shared understanding
- Examples
  - Average accuracy of each team member's estimates of information needed by other team members
  - Best situation assessment of any individual in a team
  - Consistency and overlap of shared understanding of problem, goals, information cues, and strategies

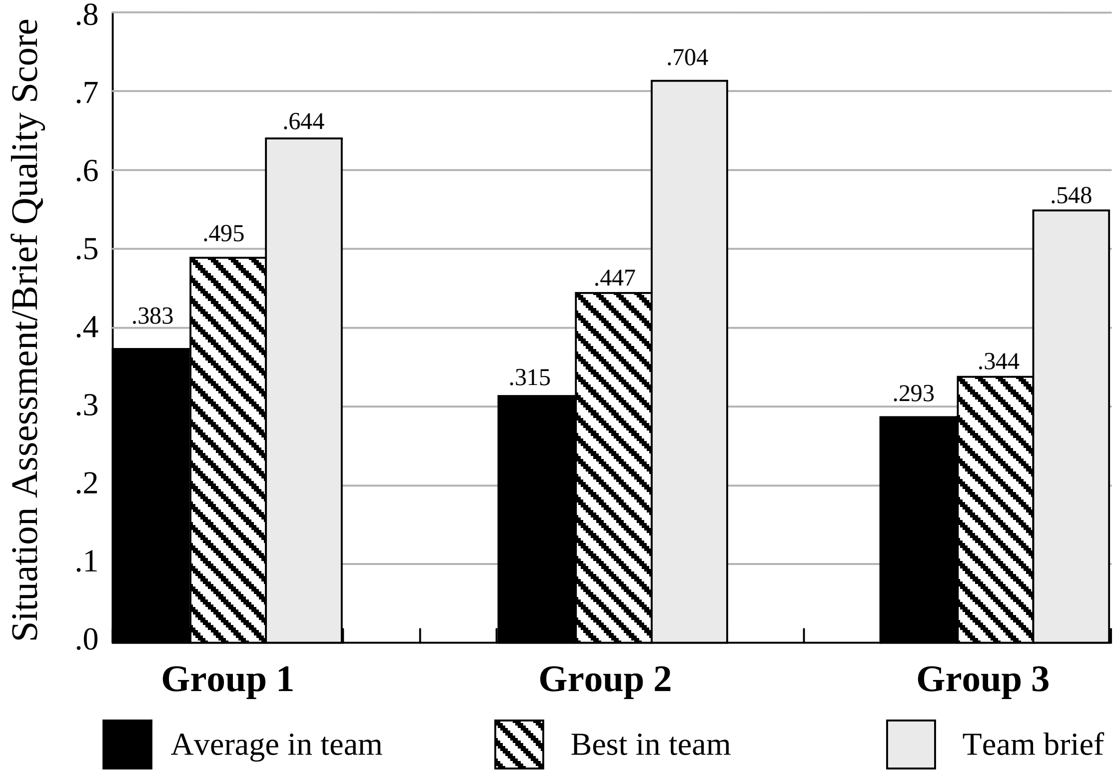
Metrics Evaluation

- Piggy-backed on two experimentation events
  - SBIR added questions to those asked by primary investigators
- The ACOA Military Utility Assessment
  - Held at PACOM May 15-20, 2001 to evaluate software utility
  - SBIR Collected data on individual and team understandings and on individual and team behaviors
- Information Superiority experiment
  - Held at JFCOM August 20-29, 2001 to compare effectiveness of alternative situation presentations
  - Collected data on individual and team understandings and on team product quality

Metrics Evaluation
Summary of Results

- Demonstrated feasibility and practicality of cognitive-oriented collaboration metrics
  - Useful data can be collected even with "fly-on-the-wall" constraints imposed by experiment organizers
  - Cognitive metrics collectable from questionnaires and inferable from behaviors
  - Team cognitive metrics practical to compute from individual cognitive metrics
  - Observed behaviors able to be linked to cognitive metrics
  - Traditional metrics for individual and team products also applicable to collaboration products

A JFCOM Result



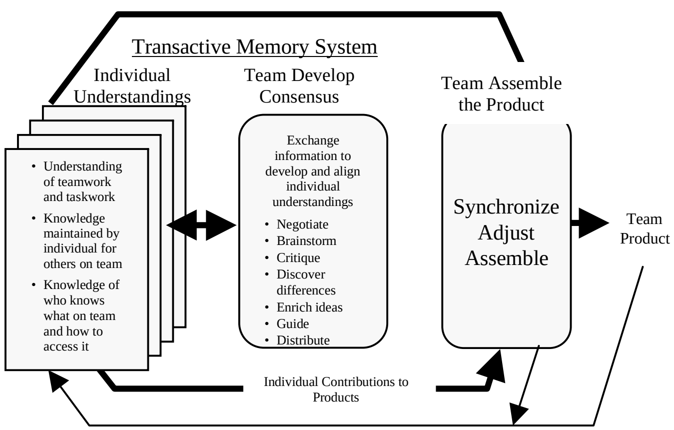
All teams always had a better briefing (team product) than either the average team understanding or the best individual understanding

Supports transactive memory model

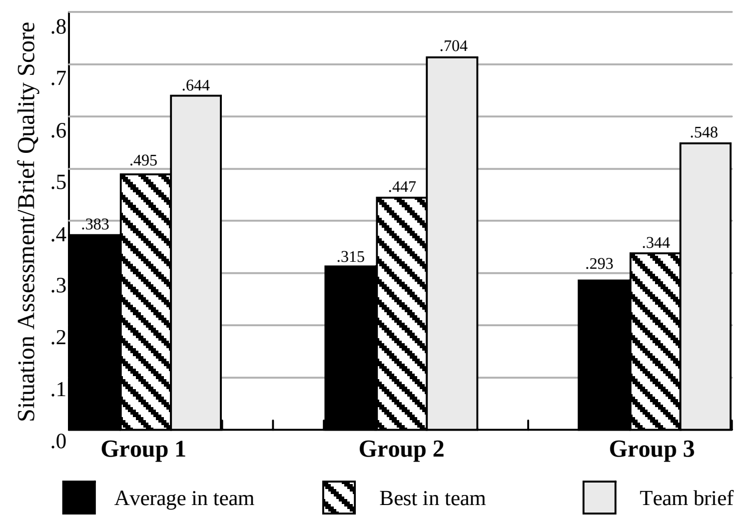
Summary

Eight categories of metrics support linking individual understandings and team product

Individual Understandings	Individual Information Interaction Support	Individual Task Performance	Individual Products
Team Understandings	Group Information Interactions	Teamwork	Team Products



Models of collaboration inform choice of metrics



Feasibility of metrics demonstrated at military experiments

Use of COBP

- Human factors
  - Experiment addressed relationships between groups and individuals
- Scenarios
  - Explicitly designed to exercise the key features expected to drive evaluation of tools
- Metrics organizations
  - Cognitive and behavioral metrics formed a hierarchy
- Use of models
  - Cognitive models of collaboration informed metrics

Backups

An ACOA MUA Result

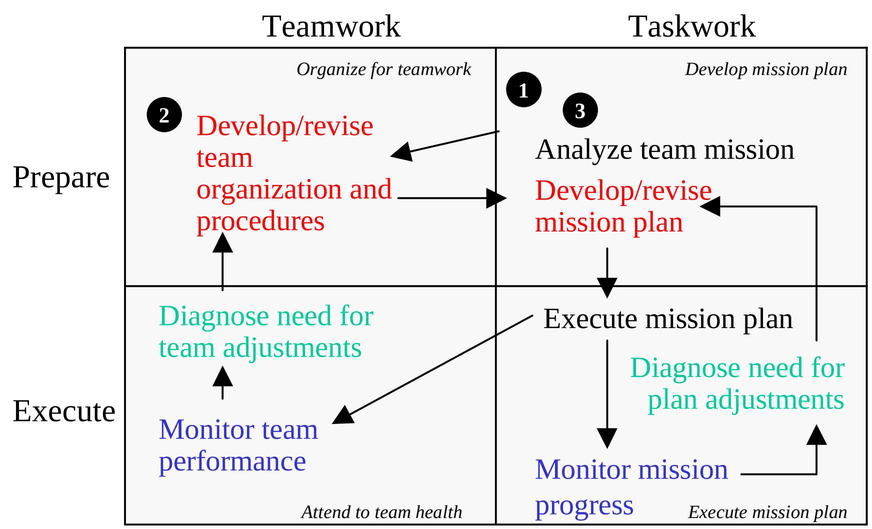
		Cognitive Causes					Non-cognitive Causes			
	Total Impacts	Tool mechanics confusion	Tool concept confusion	Task/process confusion	Low awareness of other's activities	Judged not worth the effort	Machine requires extra effort	Waiting	synchronization	Artificiality of environment
Task Impacts										
Low task quality	2		1							1
Not completed	5	1	1	2		1				
Cursory	3					3				
Delayed	4	1					2		1	
Out of order	2			2						
Increased workload	24	11	1	2	2		7			
Activity level										
low										
Engagement level										
low										
Total of Causes	41	14	3	6	2	4	9	0	1	1

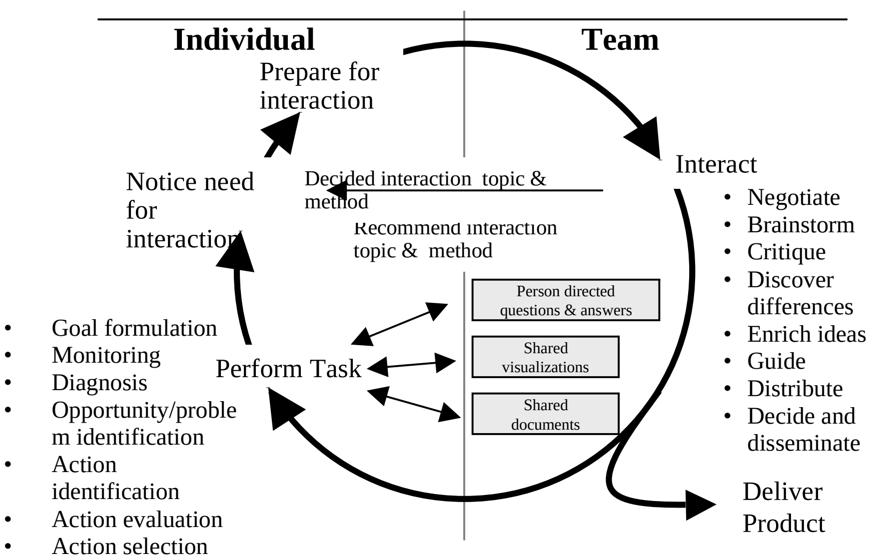
The largest negative impact of ACOA collaboration tools was increased planner workload

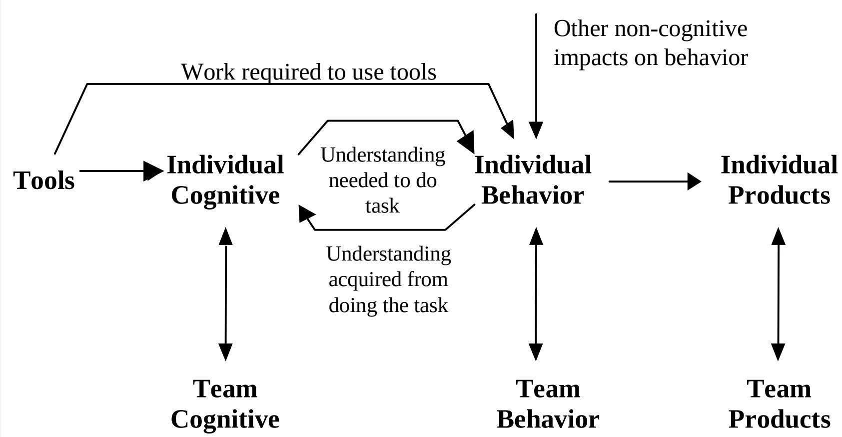
The largest contributor to increased workload was confusion on how to use tools

Complementary Collaboration Models

	Teamwork Team organization and maintenance	Taskwork Accomplishing mission goals
Prepare	Organize for teamwork • Agree on goals • Identify tasks • Assign roles • Develop schedule • Identify interaction criteria and methods	Develop mission plan • Analyze mission • Identify tasks • Allocate tasks • Develop schedule • Assign resources • Identify constraints • Develop contingencies
Execute	Attend to team health • Monitor team processes • Cue and alert to possible problems • Diagnose nature of team problem • Reengage “organize for teamwork”	Execute mission plan • Monitor • Assess situation • Decide on needed plan adjustment • Issue directives • Execute / develop products







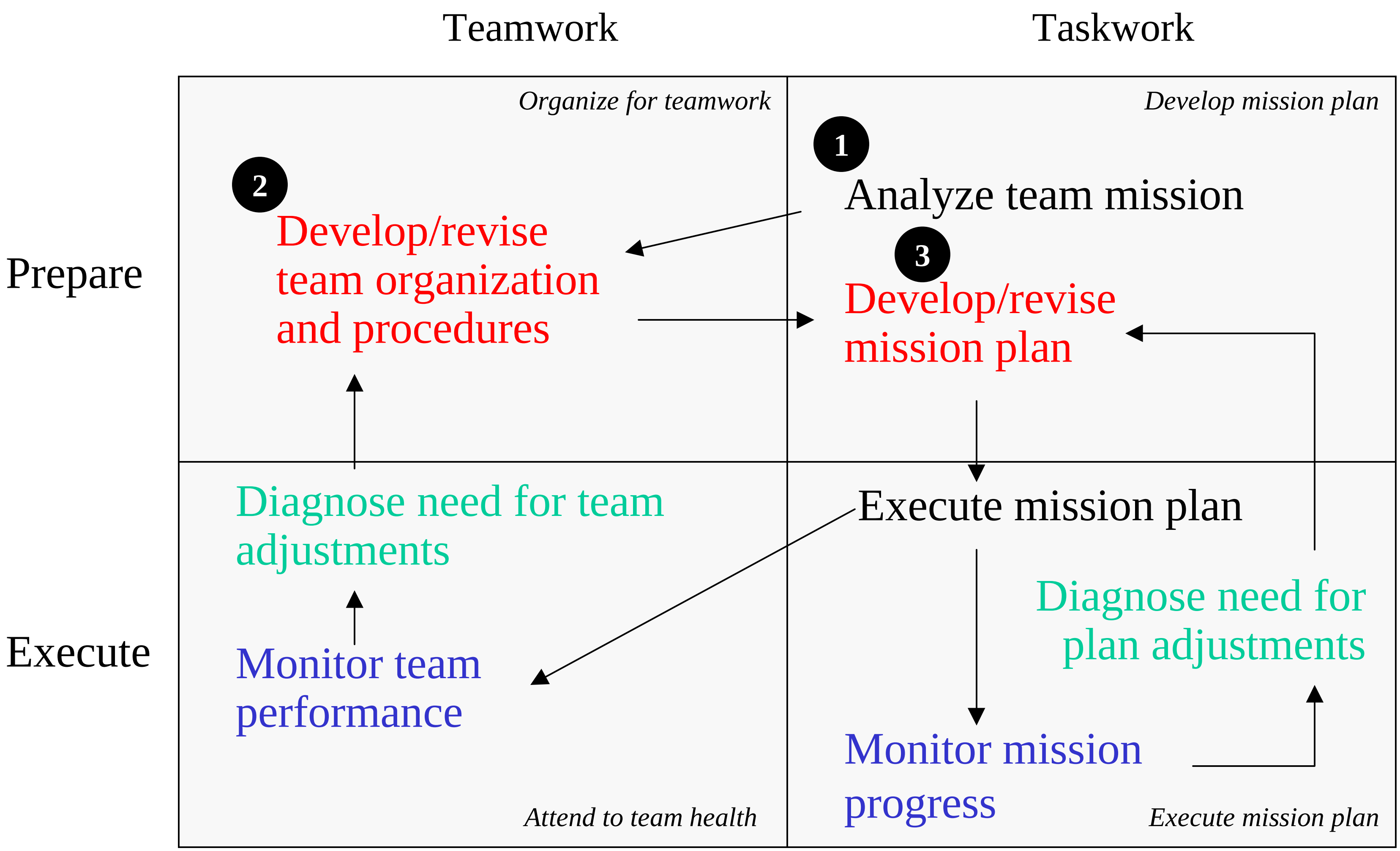
Need for Complementary Collaboration Models

- No model can likely address all aspects of collaboration
- However, a set of complementary models, each addressing different aspects of the same underlying process, can
- Five models of value to the SBIR are
  - **Teamwork and Taskwork**: all teams engage in activities to develop tasked product and to maintain team health
  - **Feedback:** teams must monitor progress and make corrections for both teamwork and taskwork
  - **Individual/Team Interplay**: many collaborative tasks call for a cycle of individual and group processes
  - **Coupling Cognition, Behavior, and Products**: team processes and products emerge from individual understandings, behaviors, and products
  - **Transactive memory**: the relationship among team member's individual understandings drives the quality of team performance

Model 1: Teamwork and Taskwork

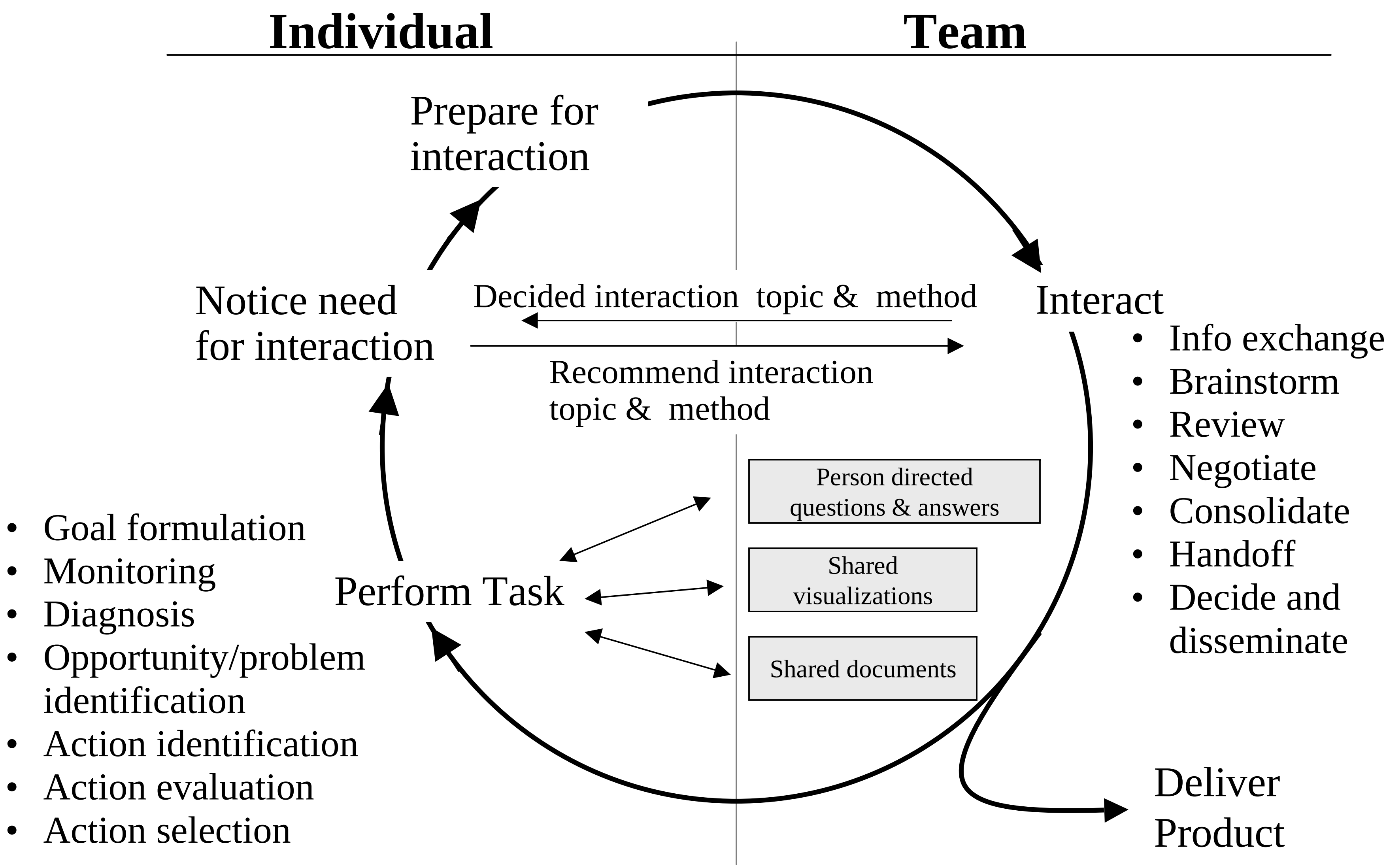
	Teamwork Team organization and maintenance	Taskwork Accomplishing mission goals
Prepare	Organize for teamwork • Agree on goals • Identify tasks • Assign roles • Develop schedule • Identify interaction criteria and methods	Develop mission plan • Analyze mission • Identify tasks • Allocate tasks • Develop schedule • Assign resources • Identify constraints • Develop contingencies
Execute	Attend to team health • Monitor team processes • Cue and alert to possible problems • Diagnose nature of team problem • Reengage “organize for teamwork”	Execute mission plan • Monitor • Assess situation • Decide on needed plan adjustment • Issue directives • Execute / develop products

Model 2
Team Planning/Execution Feedback

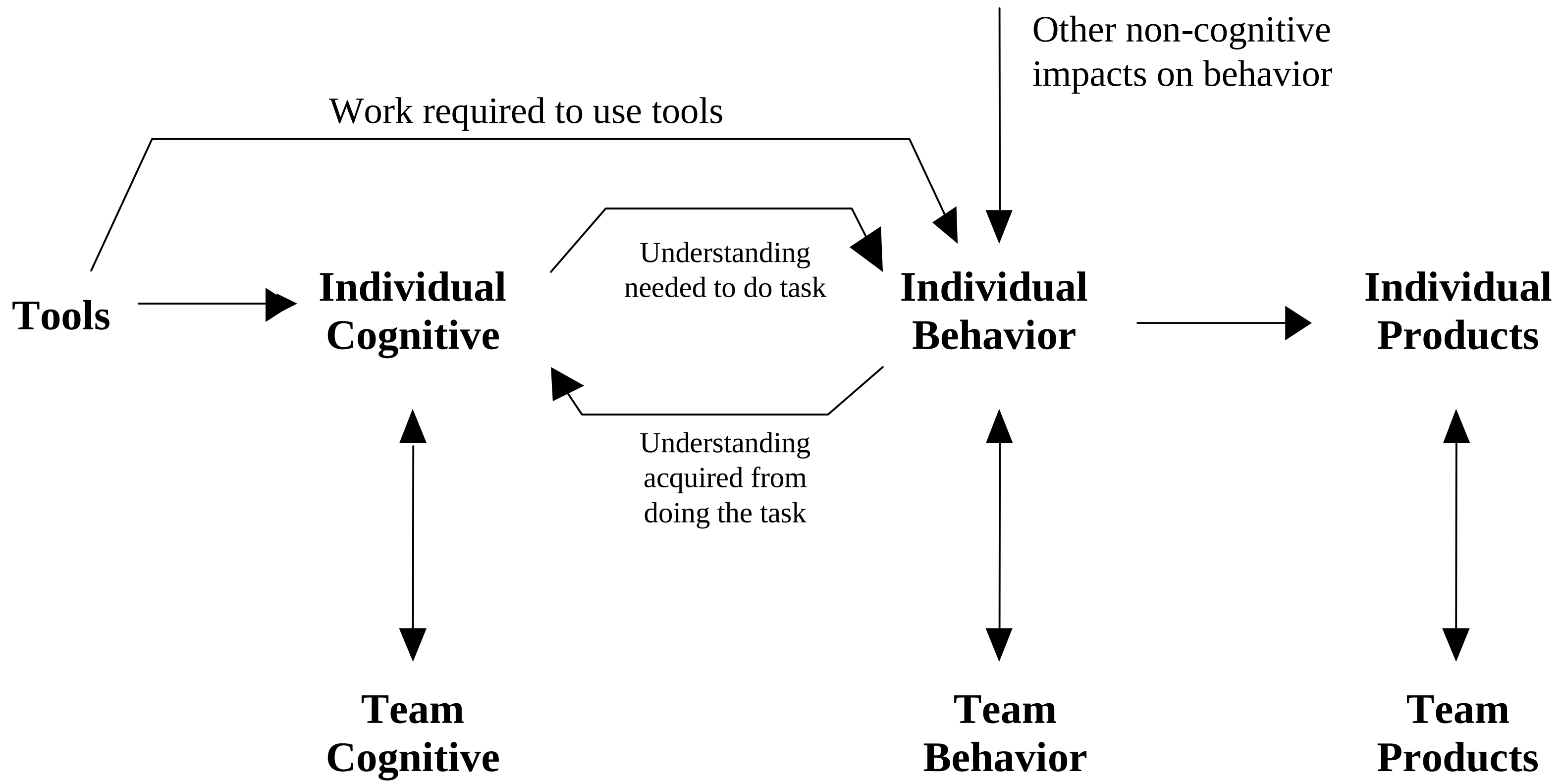


All processes may be accomplished through an interplay of individual and team work

Model 3: Individual/Team Interplay



Model 4
Coupling Cognition, Behavior, and Products



Common Ground

- What each collaboration participant assumes about each other in order to have effective interactions
- Includes each team member's assumptions about other team members'
  - Goals; e.g., where they're coming from
  - Skills, expertise, and information, to include knowledge about the external situation
  - Status, to include workload, fatigue, distraction, level of engagement
  - Degree of commitment and buy-in
  - Cognitive strategies and approach to problem solving